

WORTACH

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 16, 2000, 02:39:03 ; Search time 2352.22 Seconds  
(without alignments)  
-2062.792 Million cell updates/sec

Title: US-09-259-389-1  
Perfect score: 1598  
Sequence: 1 atggaacagacgtattga.....taggcgttgcattccaagg 1598

Scoring table: IDENTITY\_NUC

Searched: 821193 seqs, -1518192014 residues

tabase : GenEmbl.\*

Word size : 0

Number of hits that pass the threshold : 1642386

- 1: gb\_ba1.\*
- 2: gb\_ba2.\*
- 3: gb\_ov.\*
- 4: gb\_pat.\*
- 5: gb\_ph.\*
- 6: gb\_pl1.\*
- 7: gb\_pl2.\*
- 8: gb\_pr1.\*
- 9: gb\_pr2.\*
- 10: gb\_pr3.\*
- 11: gb\_ro.\*
- 12: gb\_ro.\*
- 13: gb\_sts.\*
- 14: gb\_sy.\*
- 15: gb\_un.\*
- 16: gb\_v1.\*
- 17: em\_fun.\*
- 18: em\_hum1.\*
- 19: em\_hum2.\*
- 20: em\_in.\*
- 21: em\_om.\*
- 22: em\_or.\*
- 23: em\_ov.\*
- 24: em\_pat.\*
- 25: em\_ph.\*
- 26: em\_pl1.\*
- 27: em\_ro.\*
- 28: em\_sts.\*
- 29: em\_sy.\*
- 30: em\_un.\*
- 31: em\_v1.\*
- 32: gb\_htg1.\*
- 33: gb\_htg2.\*
- 34: gb\_in1.\*
- 35: gb\_in2.\*
- 36: em\_ba1.\*
- 37: em\_ba2.\*
- 38: em\_hum3.\*
- 39: em\_hum4.\*
- 40: gb\_pr4.\*
- 41: gb\_htg3.\*
- 42: gb\_htg4.\*
- 43: gb\_htg5.\*
- 44: gb\_htg6.\*
- 45: gb\_htg7.\*
- 46: em\_htg1.\*
- 47: em\_htg2.\*
- 48: em\_htg3.\*
- 49: em\_hum5.\*

50: gb\_pl3.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1578.8	98.8	1847	12	AF044257	AF044257 Mus muscu
2	1479.6	92.6	2490	40	AF130863	AF130863 Homo sapi
3	557.2	34.9	986	4	AF024439	AF024439 Xenopus l
4	404.2	25.3	1788	10	S80876	S80876 IKAROS-hi1
5	398	24.9	1386	5	AR049700	AR049700 Sequence
6	395.8	24.8	1557	4	GGIKTRF	Y11833 G.gallus mR
7	392	24.5	3629	11	HSU040462	U040462 Human IKar
8	376	23.5	2688	12	AB017615	AB017615 Mus muscu
9	364	22.8	2049	5	AR049702	AR049702 Sequence
10	336	21.0	2496	4	OMU92201	U92201 Oncorhynch
11	325	20.3	1521	12	AF001293	AF001293 Mus muscu
12	316.8	19.8	2098	40	AF129512	AF129512 Homo sapi
13	293.6	18.4	1485	4	GGA005933	AJ005933 Gallus ga
14	286.6	17.9	1550	12	MUSIKAROS	I03547 Mouse IKaro
15	285.4	17.9	3923	4	AF186359	AF186359 Ambystoma
16	285	17.8	1788	5	AR049699	AR049699 Sequence
17	272.2	17.0	2183	4	OMU92200	U92200 Oncorhynch
18	239.8	15.0	2301	4	OMU922198	U922198 Oncorhynch
19	235.6	14.7	2309	4	AF092175	AF092175 Danio rer
20	221.6	13.9	684	12	S74708	S74708 IKaros/LyF-
21	192.8	12.1	1004	5	AR049705	AR049705 Sequence
22	190.2	11.9	1296	5	AR049701	AR049701 Sequence
23	177.6	11.1	2079	4	OMU92199	U92199 Oncorhynch
24	174.4	10.9	415	4	XLU92202	U92202 Xenopus lae
25	156.6	9.8	126836	44	AC013736	AC013736 Homo sapi
26	153.2	9.6	1170	5	AR049703	AR049703 Sequence
27	146.6	9.2	1128	5	AR049704	AR049704 Sequence
28	105.2	6.6	148640	4	AF056116	AF056116 Fugu rubr
29	102.6	6.4	41081	41	AC010286	AC010286 Homo sapi
30	102.6	6.4	80397	41	AC011502	AC011502 Homo sapi
31	102.6	6.4	90908	42	AC008463	AC008463 Homo sapi
32	100.8	6.3	118767	9	HS4PTEL	Z95704 Human DNA S
33	100.8	6.3	2088	9	HS4PZNF1	Z956138 H.septiens t
34	100	6.3	3839	9	HUMRUPZN	L11672 Human Krupp
35	97.8	6.1	2582	9	HUMZIFI	L32163 Homo sapien
36	97.8	6.1	4208	11	AF011573	AF011573 Homo sapi
37	97.8	6.1	2582	13	G28565	G28565 human STS S
38	97.8	6.1	40404	41	AC011538	AC011538 Homo sapi
39	97	6.1	236008	42	AC012616	AC012616 Homo sapi
40	96.2	6.0	3003	10	HSZNF43	X59244 Human ZNF43
41	94.6	5.9	2873	9	HUMHPLK	M55422 Human Kruep
42	94.6	5.9	2873	13	G28705	G28705 SWSS3976 Er
43	94.6	5.9	2873	13	HUMSWS1269	G18281 human chrom
44	94.4	5.9	2985	11	HUMZNFING	L15309 Human zinc
45	93.8	5.9	1434	12	MMNZFP	X52533 Mouse MRNA

ALIGNMENTS

RESULT 1  
AF044257 AF044257 1847 bp mRNA ROD 04-FEB-1998  
LOCUS Mus musculus multi-zinc finger protein helios mRNA, complete cds.  
DEFINITION  
ACCESSION AF044257  
VERSION AF044257.1 GI:2829276  
KEYWORDS  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 1847)

AUTHORS Hahn, K., Cobb, B.S., McCarty, A.S., Brown, K.E., Kling, C.A., Lee, R.,  
Akashi, K., Weissmann, I.L., Fisher, A.G. and Smal, S.T.  
TITLE Direct Submission  
Submitted (22-JAN-1998) HMI/M&I, UCLA, 675 Circle Drive South, Box  
951662, Los Angeles, CA 90095-1662, USA  
FEATURES Location/Qualifiers  
1. 1847  
source /organism="Mus musculus"  
db\_xref="taxon:10090"  
178..1758  
CDS /note="Ikaro's family member"  
/codon\_start=1  
/product="multi-zinc finger protein helios"  
/protein\_id="AAC00513.1"  
/db\_xref="GI:2829277"  
/translation="METDAIGYITCDNELSPGEHANMAIDLTSTTPNGOHASPSHM  
TSINSVKLEQSDCEDCDROPLSGDEIRGHDEGSLEELPIESSEVAQNRKVDLOGE  
GGIRLPLNGKLDYGVCMVIGPNVLVHKRSHGERPFCNOCGASFTQKGLRHFK  
LHSGEPFKPCFVSACRRDALGHLRTHSVGPKHCNYCKRSYKQSSLEEKERC  
HNTLQNVMEAAQVMVSHVPPMEDCKEQEIMDNISLVPPERPAVLEKLTANMGR  
KSTTPQKEVGEKLMRFSDYDFHDMNTYEKAEALQSHMDQAINNAITLYGAELH  
PLMGAPSTAEVAPVISAYSOVYHPNRIERTSRETSDSHENNDGPTSLRPKSR  
POBERASPSNCLSDTSSESHDDROSVQGNPALNPKRQSPAYMKEDVKALDAP  
KGSILKDYKVENGEQIATRAFKCEHCRVLFIDHVMYTIHMGCHGYRDPLECNIGYS  
ODRYESSHVIRGEHTF"

misc\_feature 517..834  
/note="encodes N-terminal zinc finger domain"  
1594..1746  
misc\_feature /note="encodes C-terminal zinc finger  
domain/hunchback-like zinc fingers"  
BASE COUNT 534 a 466 c 445 g 402 t  
ORIGIN

Query Match 98.88; Score 1578.8; DB 12; Length 1847;  
Best Local Similarity 99.28; Pred. No. 0;  
Matches 1586; Conservative 0; Mismatches 12; Indels 0; Gaps 0;  
Qy 1 atggaaacagcgtattgaggtatataacatgtgacaatgagctttccaccggaagg 60  
Db 178 ATGGAAACAGCAATTGATGGCTATATAACATGTGACATGAGCTTCCACCGAAGG 237  
Qy 61 gaacagccaatattgcccattgacctcaactcaagcagcccaatggagcagcgcctcg 120  
Db 238 GAACAGCCCAATATGGCAATTGACCTCACCTCAAGCACACCCCAATGGACAGCAGCCTCG 297  
Qy 121 ccaagtccatgacacagcaccaattctgtaagctggaatgcagagtgtgagagtggt 180  
Db 298 CCAAGTCACATGACAAACACCAATTCGTAAAGCTGGAATGCAAGTGTGATGAAGTGT 357  
Qy 181 gacagcagccctgagccgtgagatgagatcaggggcccacgatgaggggagcagccta 240  
Db 358 GACAGCAGCCCTGAGCCGTGAGGATGAGATCAGGGGCCACGATGAGGGAGGAGCGCTA 417  
Qy 241 gaagaaccccttaattgagcagcagcagcagcagcagcagcagcagcagcagcagc 300  
Db 418 GAAGAACCCTTAATTGAGCAGCAGCGAGGTGGCCGACACAGGAAAGTCAGGACCTCAA 477  
Qy 301 ggcagggagggaatccggcttccgaatggtaaaactggaactggaactggtgagcaggtt 360  
Db 478 GCGAGGGAGGAATCCGGCTTCCGAATGGTAAACTGAAATGTGAGCTGTGGCATGGTT 537  
Qy 361 tgcattggggcccaatgtcttatggtacataaaaggagtcacactggtgagcggcccttc 420  
Db 538 TGCATTGGGCCCAATGTGCTTATGTTACATAAAAAGGAGTCACATGTGAGCGGCCCTTC 597  
Qy 421 caactgtaaacagtgaggagctctttaccagaaagggcaactcttgagacacataaag 480  
Db 598 CACTGTAAACAGTGGGAGCTCTTTTACCAGAAAGGCAACCTTCTGAGCACATTAAG 657  
Qy 481 tcaactctggagagaagcccttcaaatgtcctttctgtatgtatgtctgttagaagaagg 540  
Db 658 TTACTCTGGAGAGAAGCCCTTCAATGTCTTCTTCTGTAGTATGCTTGTAGAAGAAG 717

Qy 541 gacgtctcacaggacacctcagaccattctgtgggtaaacctcacaagtgaactac 600  
Db 718 GAGCTCTCACAGACACACTCAGAGCCACTCTGTGGGTAACTCAAGCTCAGAGTGAAC 777  
Qy 601 tgtggccgaagctacaagcagcagcagcagcagcagcagcagcagcagcagcagc 660  
Db 778 TGTGGCCGAAGCTACAAGCAGCGAGCTCACTGGAGAAACACAGGAACGCTGTCAAC 837  
Qy 661 tatctcagaatgtcagcagcagcagcagcagcagcagcagcagcagcagcagc 720  
Db 838 TATCTCAGAAATGTGAGCTGAGGCTGCCGGCAGGCTGATGATGATGATGATGATG 897  
Qy 721 atggaagattgtaaggaacaagagccttatcatggaacaacattattctctgtgctttt 780  
Db 898 ATGGAAGATTGTAAGGAACAAGAGCCTATCATGCACACAATATTTCTGTGTGCTTTT 957  
Qy 781 gagagacctgctgtcatagagaagcctcagcgcaaatatgggaagcgcaaaagctccact 840  
Db 958 GAGAGACCTGCTGTCTATAGAGAAGCTCAGCGCAAAATATGGGAAGCGCAAAAGCTCCACT 1017  
Qy 841 cctcagaagttgtgtgggggaaagccttatcgattcagctaccocagatattcatttggat 900  
Db 1018 CCTCAGAAGTTGTGGGGGAAAGCCTTATCGGATTGAGCTACCCAGATATTTCAATTTGAT 1077  
Qy 901 atgaacttaacatatgagaagcagcagcagcagcagcagcagcagcagcagcagc 960  
Db 1078 ATGACTTTAATATATGAGAGAGGCTGAGCTGATGCTGCTCATATGATGACCAAGCC 1137  
Qy 961 atcaacaatgcaatcacctacacttggcgtgagcgttcccttcccttcccttcccttcc 1020  
Db 1138 ATCAACAATGCAATCACCTTACCTTGGAGCTGAGGCGCTTACCCCTCTGATGACGATGCA 1197  
Qy 1021 ccaagcacatgctgaggtggcccgcttataagctcagcttattctcaggtctatcat 1080  
Db 1198 CCAAGCACATCGCTGAGGTGGCCCGAGTTATAGCTCAGCTTATTTCTCAGTCTATCAT 1257  
Qy 1081 ccaacaaggatgaaagaccattagcagggaaacatctgtagtcacagaaacaaacatg 1140  
Db 1258 CCAACAAGGATGAAAGACCCATTAGCAGGGAAACATCTGTAGTCAGCAAAACACATG 1317  
Qy 1141 gatggcccatctctctcatcagcacaagagtcgaccccgaggaagagagcgtcgcgc 1200  
Db 1318 GATGGCCCATCTCTCATCAGACCAAGAGTCGACCCCGAGGAAAGAGAGAGAGAGAG 1377  
Qy 1201 agcaatagctgcctgattcttactgactcagcaggaagtagcagcagcagcagcagc 1260  
Db 1378 AGCAATAGCTGCCTCGATTCTACTGACTCAGAAAGTAGCCATGATGACCGCAGTCTCTAC 1437  
Qy 1261 caagaaacccctgcttataatcccaagaggaacaaagcccagcttatacagagagagat 1320  
Db 1438 CAAGAAACCCCTGCTTAAATCCCAAGAGGAAACAAAGCCAGCTTACATGAAGAGGAT 1497  
Qy 1321 gtaagccttggatgctacacattcacaagggcccaagggcctctctgaaagacattataaggtt 1380  
Db 1498 GTCAAGCCTTTGGATGTCTACCAAGGCCCGCCAGGGCTCTCTGAAGGACATCTATAAGGTT 1557  
Qy 1381 ttcaatggagaaggaacagcagcagcagcagcagcagcagcagcagcagcagcagc 1440  
Db 1558 TTCAATGGAGAAGAGACAGATAAGGCGCTTCAAGTGTGAGCACTGCCGAGTCTTTT 1617  
Qy 1441 ctagaacattgcatgacacattcacaatgggttggccatggcttaccgggagccactgaa 1500  
Db 1618 CTAGACCATGTCATGTACACCATTCACATGGGTGGCATGGCTTACCGGAGCCACTGAA 1677  
Qy 1501 tgcacatctgtgcttacagaagccagcagcagcagcagcagcagcagcagcagcagc 1560  
Db 1678 TGCAACATCTGTGGCTACAGAAGCCAGGACCGCTACGAATTTTCATCACATTTGTCGA 1737  
Qy 1561 gggcagacacattcactcagcagcagcagcagcagcagcagcagcagcagcagc 1598  
Db 1738 GGGAGGACACACATTCACCTAGGCGCTTTTTCATTTCCAAAG 1775





GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 16, 2000, 03:23:33 ; Search time 2352.22 Seconds  
(without alignments)  
-1962.105 Million cell updates/sec

Title: US-09-259-389-3  
Perfect score: 1520  
Sequence: 1 atggaacacagcatttga.....tagcggttgcatccaagg 1520

Scoring table: IDENTITY\_NUC

Searched: 821193 seqs, -1518192014 residues

Database: GenEmbl.\*

Word size: 0

Number of hits that pass the threshold : 1642386

- 1: gb\_bal.\*
- 2: gb\_ba2.\*
- 3: gb\_om.\*
- 4: gb\_ov.\*
- 5: gb\_pat.\*
- 6: gb\_ph.\*
- 7: gb\_pli.\*
- 8: gb\_pl2.\*
- 9: gb\_pri.\*
- 10: gb\_pr2.\*
- 11: gb\_pr3.\*
- 12: gb\_ro.\*
- 13: gb\_sts.\*
- 14: gb\_sy.\*
- 15: gb\_un.\*
- 16: gb\_vi.\*
- 17: em\_fun.\*
- 18: em\_hum1.\*
- 19: em\_hum2.\*
- 20: em\_in.\*
- 21: em\_om.\*
- 22: em\_or.\*
- 23: em\_ov.\*
- 24: em\_pat.\*
- 25: em\_ph.\*
- 26: em\_pl.\*
- 27: em\_ro.\*
- 28: em\_sts.\*
- 29: em\_sy.\*
- 30: em\_un.\*
- 31: em\_vi.\*
- 32: gb\_htg1.\*
- 33: gb\_htg2.\*
- 34: gb\_in1.\*
- 35: gb\_in2.\*
- 36: em\_bal.\*
- 37: em\_ba2.\*
- 38: em\_hum3.\*
- 39: em\_hum4.\*
- 40: gb\_pr4.\*
- 41: gb\_htg3.\*
- 42: gb\_htg4.\*
- 43: gb\_htg5.\*
- 44: gb\_htg6.\*
- 45: gb\_htg7.\*
- 46: em\_hcg1.\*
- 47: em\_hcg2.\*
- 48: em\_hcg3.\*
- 49: em\_hum5.\*

50: gb\_pl3.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1412.8	92.9	1847	12	AF044257 Mus muscu
2	1313.6	86.4	2490	40	AF130863 Homo sapi
3	557.2	36.7	986	4	AF024439 Xenopus l
4	330.4	21.7	1557	4	Y11833 G.gallus mr
5	329.6	21.7	1386	5	AR049700 Sequence
6	329.6	21.7	1788	10	S80876 IKAROS-hiK1
7	317.4	20.9	3629	11	U40462 Human Ikaro
8	289.6	19.1	1521	12	AF001293 Mus muscu
9	286.4	18.8	2049	5	AR049702 Sequence
10	285.8	18.8	1550	12	MUSIKAROS
11	285.4	18.8	3923	4	AF186359 Ambystoma
12	284.2	18.7	1788	5	AR049699 Sequence
13	283	18.6	2688	12	AB017615 Mus muscu
14	279.2	18.4	2098	40	AF129512 Homo sapi
15	273.6	18.0	2496	4	OMU92201 Oncorhynch
16	272.2	17.9	2183	4	OMU92200 Oncorhynch
17	247	16.2	1485	4	GGA005933 Gallus ga
18	216.4	14.2	2309	4	AF092175 Danio rer
19	192.8	12.7	1004	5	AR049705 Sequence
20	179	11.8	2301	4	OMU92198 Oncorhynch
21	177.6	11.7	2079	4	OMU92199 Oncorhynch
22	174.4	11.5	415	4	XLU92202 Xenopus lae
23	164	10.8	684	12	S74708 IKAROS/LyF-
24	157.4	10.4	126836	44	AC013736 Homo sapi
25	153.2	10.1	1170	5	AR049703 Sequence
26	146.6	9.6	1296	5	AR049701 Sequence
27	146.6	9.6	1128	5	AR049704 Sequence
28	105.2	6.9	148640	4	AF056116 Fugu rubr
29	81.4	5.4	1940	12	RNU67082 Rattus norv
30	81.2	5.3	3068	12	AF111101 Mus muscu
31	80.2	5.3	236008	42	AC012616 Homo sapi
32	80	5.3	1434	12	MMNZFP Mouse mRNA
33	80	5.3	3492	12	MMZFP37P1
34	80	5.3	1949	12	MMZFP
35	78.4	5.2	153875	11	AC003682 Homo sapi
36	77.4	5.1	3839	9	HUNKRUP2N
37	76.6	5.0	118767	9	HS4PTEL
38	76.6	5.0	2088	9	HS4PZNF1
39	75.8	5.0	147193	42	AC010467 Homo sapi
40	75.4	5.0	1661	12	MMA224763 Mus muscu
41	75.2	4.9	2971	12	AF072439 Rattus no
42	75.2	4.9	206121	44	AC016769 Homo sapi
43	75	4.9	2582	9	HUMZIFI
44	74	4.9	3003	10	HSZNF43
45	75	4.9	4208	11	AF011573 Homo sapi

ALIGNMENTS

RESULT 1  
AF044257 AF044257 1847 bp mRNA ROD 04-FEB-1998  
LOCUS Mus musculus multi-zinc finger protein helios mRNA, complete cds.  
DEFINITION  
ACCESSION AF044257  
VERSION AF044257.1 GI:2829276  
KEYWORDS house mouse.  
SOURCE Mus musculus  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 1847)



```

RESULT 2
AF130863 2490 bp mRNA PRI 22-NOV-1999
LOCUS Homo sapiens zinc finger DNA binding protein Helios mRNA, complete cds
DEFINITION
ACCESSION AF130863
VERSION AF130863.1 GI:6457257
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2490)
AUTHORS Hosokawa,Y., Maeda,Y. and Seto,M.
TITLE Human Helios, an ikaros-related zinc finger DNA binding protein:
cDNA cloning and tissue expression pattern
JOURNAL Immunogenetics 50 (1-2), 106-108 (1999)
MEDLINE 20009537
PUBMED 10640811
AUTHORS Hosokawa,Y., Maeda,Y. and Seto,M.
TITLE Direct Submission
JOURNAL Submitted (24-FEB-1999) Lab. of Chemotherapy, Aichi Cancer Center
Research Institute, 1-1 Kanakoden, Chikusa-ku, Nagoya, Aichi
464-8661, Japan
FEATURES
source
1..2490
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="2"
/map="2q13.1"
170..1750
/codon_start=1
/product="zinc finger DNA binding protein Helios"
/protein_id="AAF09441.1"
/db_xref="GI:6457258"
/translation="METDAIDGYITCDNELSPEREHSNMAIDLTSSPNQHASPSHM
TSTNKLQMDQDECDKPLSRENEIRGRHDEGSLSEPLIESSEVADNKNKVOELGE
GGTRLPNGKLKCDVGMVCLGNVLVHVKRSHGPRPHQCQASTQKGNLRLRIK
LHSEKPEKPCSVACRRDALTLHLRHSVGHKNCYGRSKRSLEHKEKRC
HNYLQNSNEARQKMSHVPPMEDCKEPEIMDNNISLVFFERPAVIEKLTGNMGRK
KSTPKFVEKLMKSYFDIHEMLNLTVEKAEMLQSHMDQAINNAITLGLAEALH
PLMQHPSTIAEAPVTSISAYSQVYHPNRIERIPIRETADSHENMDGPISLRPKSR
PQERASPSNCLDSTSESHDHQSYQHPALNPKQSPAYMKEDVKAIDTTKAP
KGLKDIYKVFNGEQQIRAFKECHRVLFIDHVMYTHMGCHGYRPLECNICGYS
ODRYEFSSHIVRGETHF"
BASE COUNT 741 a 553 c 571 g 625 t
TGIN
Query Match 86.4%; Score 1313.6; DB 40; Length 2490;
Best Local Similarity 90.5%; Pred. No. 0;
Matches 1446; Conservative 0; Mismatches 74; Indels 78; Gaps 1;
QY 1 atggaacacacgctattagctatataacatgtacaaatgagctttcaccccaagg 60
Db 170 ATGGAACACAGCGCTATTGATGGCTATATAACGTGTGACAAATGAGCTTTACCCGAAAG 229
QY 61 gaacacgccaatattggccattgacctcactcaagcagcgcaccaatggacagcacgctcg 120
Db 230 GAGCACTCCAATATGGCAATTGACCTACCTCAAGCACACCCCAATGGACAGCATGCTCA 289
QY 121 ccaagtacacatgacagcaacaattctgtaagctggaaatgacagatgagagatgt 180
Db 290 CCAAGTACATGACAGCAACAATTCAGTAAAGCTAGAAATGACAGAGTATGAAGAGTGT 349
QY 181 gacaggcagccctgagcgtgagatgagatcaggggcccacgacgatgagggagcagccta 240
Db 350 GACAGGAACCCCTGAGCCGTGAAATTCAGATCAGGGGCCATGATGAGGTAGCAGCTA 409
QY 241 gaacaccccttaattgagacgagctggccgacaaacgaagaagctccaggaacctta 300
Db 410 GAAGAACCCCTTAATTGAGACGACGAGGTGGCTGACCAACAGGAAGAGTCCAGAGGCTCAA 469

```

```

QY 301 ggcgaggagggaatcggcttcgaa----- 327
Db 470 GCGAGGAGGAATCCGCTTCCGAATGGTAAACTGAAATGTGACGTCTGTGGCAGTTT 529
QY 327 -----*tggtgagcggcccttc 342
Db 530 TGCATTGGGCCCAATGTCTTATGGTACATAAAAGGAGTCACACTGGTGAACGCCCTTC 589
QY 343 cactgaacacgctcgagctcttttaccagaagggcaaccttctgagacacataaag 402
Db 590 CATGTAAACAGTGTGGAGCTTCTTTACTCAGAAGGGCACTTCTTGAGACATAAAG 649
QY 403 ttacactctggagagaagcccttcaaatgtcctttctgtagctatgctttagaagaag 462
Db 650 TTACTCTGGAGAGAAGCCGTTCAAAATGTCTTCTGTAGTACGCCCTGTAGAAGAAG 709
QY 463 gacgctctcacaggacacctcagagccattctgtgggtaaacctcacaagtgaactac 522
Db 710 GAGCCCTCACAGGACACTCAGGACCCATTCTGGGTAAACCTCACAGTGCACACTAC 769
QY 523 tgtggccgaagtacaagcagcgcaogtcaactggaggacaacaaggacgcgtgtcaaac 582
Db 770 TGTGGAGGAAGCTACAGCAGCGCAGTCTACTGGAGGAGCACAAGAACGCTGCCAAC 829
QY 583 tatctccagaatgtcagctgaggtgcgaggcaggtcaggtcaggtcaggtcaggtc 642
Db 830 TATCTCCAGATGTGACGATGGAGGCTGTGGCAGGTCTATGATGATGATGATGATG 889
QY 643 atggaagattgttaaggaaagagccttctatcttggaacaacaattctctgtgtgtct 702
Db 890 ATGGAAGATTGTAAAGAACAGAGCGCTTATTATGGACAACAATATTTCTGTGTCTTT 949
QY 703 gagagacctgtctcatagagaagctcagcgaataatggaagcgaacaaagcctccat 762
Db 950 GAGAGACCTGTCTCATAGAGAGCTCAGCGGGAATATGGGAACATGTAAGACTCCAT 1009
QY 763 cctcagaagttgtgggggaaagcttatgcattcagctaccaccagatatctattgtat 822
Db 1010 CCACAAAAGTTGTGGGGGAAAAGCTCATGCGATTAGCTACCCAGATATTCCTTTGAT 1069
QY 823 atgaactaaacatatagagaaggggtgcagctgagctcagctcagctcagctcagctc 882
Db 1070 ATGAACCTTAACATATGAGAGGAGGCTGAGCTGATGATGATGATGATGATGATGAT 1129
QY 883 atcaacaatgcaatcactcacttggagctgagctgagcctcactcactcactcactc 942
Db 1130 ATCAACAATGCAATCACCCTACCTTGGAGCTGAGGCCCTTCACCCCTCTGATGACGAC 1189
QY 943 ccaagcacaatcgtgaggtggcccgcttaagctcagctcagctcagctcagctcagctc 1002
Db 1190 CCAAGCACAATCGCTGAAGTGGCCCGAGTTTAAAGCTAGCTTATTTCTCAGGTCTAT 1249
QY 1003 ccaacacgagtagaagacccattagcagggaacacatctgacgaggaacacacatg 1062
Db 1250 CCAAAATAGATAGAAAGACCCATTAGCAGGGAACCTGTGATAGTATGATGATGATGAT 1309
QY 1063 gatggcccatctctcatcagacaaagagctcagcccgaggaagagagagagagagag 1122
Db 1310 GATGGCCCATCTCTCTCATCAGACCAAGAGTCCACCCAGGAAGAGAGGCTCTCCC 1369
QY 1123 agcaatagctgcctgatttacttactcagaaatgacatgacgagccagctcagctc 1182
Db 1370 AGCAATAGTGCCTGGATTCCACTGACTCAGAAAGAGCCATGATGATGATGATGATGAT 1429
QY 1183 caaggaaacctctgcttaaatccaaagagaaacacacagctcagctcagctcagctcag 1242
Db 1430 CAAGGACACCTCGCTTAAATCCCAAGAGAAACAAACCCAGCTTACATGAAGAGGAT 1489
QY 1243 gtcaagggtttgagatgtaaccaggcccccaagggtctcttgaaggagacatctataag 1302
Db 1490 GTCAAAGCTTTGGTACTACCAAGGCTCTTAAGGGCTCTCTGAAGGACATCTACAAG 1549

```





GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 16, 2000, 03:24:16 ; Search time 2352.22 Seconds  
(without alignments)  
-2487.485 Million cell updates/sec

Title: US-09-259-389-5  
Perfect score: 1927  
Sequence: 1 gccggggcagtcgcattgc.....gcagtcattgtctgtctgt 1927

Scoring table: IDENTITY\_NUC

Searched: 821193 seqs, -1518192014 residues

abase : GenEmbl:\*

Word size : 0

Number of hits that pass the threshold : 1642386

- 1: gb\_ba1:\*
- 2: gb\_ba2:\*
- 3: gb\_om:\*
- 4: gb\_ov:\*
- 5: gb\_pat:\*
- 6: gb\_ph:\*
- 7: gb\_pl1:\*
- 8: gb\_pl2:\*
- 9: gb\_pl3:\*
- 10: gb\_pr2:\*
- 11: gb\_pr3:\*
- 12: gb\_ro:\*
- 13: gb\_sts:\*
- 14: gb\_sy:\*
- 15: gb\_un:\*
- 16: gb\_v1:\*
- 17: em\_fun:\*
- 18: em\_hum1:\*
- 19: em\_hum2:\*
- 20: em\_in:\*
- 21: em\_om:\*
- 22: em\_or:\*
- 23: em\_ov:\*
- 24: em\_pat:\*
- 25: em\_ph:\*
- 26: em\_pl:\*
- 27: em\_ro:\*
- 28: em\_sts:\*
- 29: em\_sy:\*
- 30: em\_un:\*
- 31: em\_v1:\*
- 32: gb\_htg1:\*
- 33: gb\_htg2:\*
- 34: gb\_in1:\*
- 35: gb\_in2:\*
- 36: em\_ba1:\*
- 37: em\_ba2:\*
- 38: em\_hum3:\*
- 39: em\_hum4:\*
- 40: gb\_pr4:\*
- 41: gb\_htg3:\*
- 42: gb\_htg4:\*
- 43: gb\_htg5:\*
- 44: gb\_htg6:\*
- 45: gb\_htg7:\*
- 46: em\_htg1:\*
- 47: em\_htg2:\*
- 48: em\_htg3:\*
- 49: em\_hum5:\*

50: gb\_pl3:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1670.8	86.7	2490	40	AF130863	AF130863 Homo sapi
2	1593.8	82.7	1847	12	AF044257	AF044257 Mus muscu
3	546	28.3	986	4	AF024439	AF024439 Xenopus l
4	401.6	20.8	1788	10	S80876	S80876 IKAROS-hik1
5	400.6	20.8	1386	5	AR049700	AR049700 Sequence
6	394.2	20.5	1557	4	GKIKTRF	Y11833 G.gallus mR
7	394.2	20.5	3629	11	HSU40462	U40462 Human Ikaro
8	381	19.8	2688	12	AB017615	AB017615 Mus muscu
9	365.8	19.0	2049	5	AR045702	AR045702 Sequence
10	339.4	17.6	1521	12	AF001293	AF001293 Mus muscu
11	337.6	17.5	2496	4	OM092201	U92201 Oncorhynch
12	331.2	17.2	2098	40	AF129512	AF129512 Homo sapi
13	298.4	15.5	1485	4	GGA005933	AJ005933 Gallus ga
14	297.2	15.4	3923	4	AF186359	AF186359 Ambystoma
15	293.6	15.2	1550	12	MUSIKAROS	L03547 Mouse Ikaro
16	292	15.2	1788	5	AR049699	AR049699 Sequence
17	273.8	14.2	2183	4	OM092200	U92200 Oncorhynch
18	230.2	11.9	2301	4	OM092198	U92198 Oncorhynch
19	228.2	11.8	2309	4	AF092175	AF092175 Danio rer
20	218.4	11.3	684	12	S74708	S74708 Ikaro/Lyf-
21	206.2	10.7	1004	5	AR049705	AR049705 Sequence
22	180.2	9.4	1296	5	AR049701	AR049701 Sequence
23	168	8.7	2079	4	OM092199	U92199 Oncorhynch
24	166.4	8.6	415	4	XL092202	U92202 Xenopus lae
25	164.6	8.5	126836	44	AC013736	AC013736 Homo sapi
26	162.8	8.4	1170	5	AR049703	AR049703 Sequence
27	153	7.9	1128	5	AR049704	AR049704 Sequence
28	108	5.6	148640	4	AF056116	AF056116 Fugu rubr
29	99.4	5.2	126836	44	AC013736	AC013736 Homo sapi
30	94.6	4.9	41081	41	AC010286	AC010286 Homo sapi
31	94.6	4.9	80397	41	AC011502	AC011502 Homo sapi
32	93	4.8	90308	42	AC008463	AC008463 Homo sapi
33	91.4	4.7	2088	9	HS4PZNF1	Z96138 H.sapiens t
34	91.4	4.7	2582	9	HUM2IFI	L32163 Homo sapien
35	91.4	4.7	4208	11	AF011573	AF011573 Homo sapi
36	91.4	4.7	2582	13	G28565	G28565 human sts s
37	90.8	4.7	1940	12	RNU67082	U67082 Rattus norv
38	90.6	4.7	5471	9	AB007886	AB007886 Homo sapi
39	90.6	4.7	118767	9	HS4PTEL	Z95704 Human DNA s
40	90.6	4.7	236008	42	AC012616	AC012616 Homo sapi
41	90.2	4.7	128779	11	HS29K1	Z98745 Human DNA s
42	90.2	4.7	187543	40	AC005678	AC005678 Homo sapi
43	90	4.7	132783	10	AC000378	AC000378 Human Chr
44	90	4.7	2394	11	AF003540	AF003540 Homo sapi
45	90	4.7	3326	12	RNU41164	U41164 Rattus norv

ALIGNMENTS

RESULT	1	AF130863	2490 bp	mrna	PRI	22-NOV-1999
LOCUS	AF130863	Homo sapiens	zinc finger DNA binding protein	Helios mRNA	complete	
DEFINITION	AF130863	cds.				
ACCESSION	AF130863					
VERSION	AF130863.1	GI:6457257				
KEYWORDS	human.					
SOURCE	human.					
ORGANISM	Homo sapiens					
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;					
	Eutheria; Primates; Catarrhini; Hominidae; Homo.					

REFERENCE 1 (bases 1 to 2490)  
 AUTHORS Hosokawa, Y., Maeda, Y. and Seto, M.  
 TITLE Human Helios, an Ikaros-related zinc finger DNA binding protein: DNA cloning and tissue expression pattern  
 JOURNAL Immunogenetics 50 (1-2), 106-108 (1999)  
 MEDLINE 20009537  
 REFERENCE 2 (bases 1 to 2490)  
 AUTHORS Hosokawa, Y., Maeda, Y. and Seto, M.  
 TITLE Direct Submission  
 JOURNAL Submitted (24-FEB-1999) Lab. of Chemotherapy, Aichi Cancer Center Research Institute, 1-1 Kanakoden, Chikusa-ku, Nagoya, Aichi 464-8681, Japan  
 FEATURES Location/Qualifiers  
 source 1..2490  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="2"  
 /map="2q13.1"  
 /length=2490  
 /codon\_start=1  
 CDS 170..1750  
 /product="zinc finger DNA binding protein Helios"  
 /protein\_id="AAF09441.1"  
 /db\_xref="GI:6457258"  
 /translation="METDAIDGYITCDNELSPERHSNMAIDLTSTPNQHASPSHM  
 TGTNSVKLEMSDECDKPLRENIHGHDEGSLERPLESEVADNRKYLQGE  
 GRIPLNGKLACDGMVCIQPNVLMVHKRSHTGRPHNQCAGSTOKNLLRIK  
 LHSQKPKPCSYACRRRDALTGLHRTSHSVGPKCNVCGSYKORSLEHKERC  
 HNYLQNSMEAGQVSHVHPMEDCKEPEIMDNNSILVFPFPAVIEKLIGNMGR  
 KSTQKVFQGLKMRFSYDPIHEDMLNLTKEKELQSHMDQINNAITVIGALH  
 PLMQHPTTIAEAPVISASQVYHPNRIEPISTRETDHSHENMDGPISLRPKSR  
 PQRAPSPNSCLDSTDSSESHDQSYQHPALNPKRKQSPAYMKEDVKALDTTKAP  
 KGLSLDIYVFNFGEGEIQIRAFKECHRVLFVDHVMYTHMGCHGYRPLECNIGYRS  
 QDRYEPFSHIVRGHTFH"  
 BASE COUNT 741 a 553 c 571 g 625 t  
 ORIGIN  
 Query Match 86.7%; Score 1670.8; DB 40; Length 2490;  
 Best Local Similarity 96.3%; Pred. No. 0;  
 Matches 1811; Conservative 0; Mismatches 22; Indels 47; Gaps 8;  
 QY 57 tccctcttctgaaacccacattgacattcctctcttga-gggttagaagaagctaga 115  
 DB 73 TCCGTTTATCTTACGCCGACATTTGTCACCCCTTTTGAGGGGTTAGAGAGCTGAGA 132  
 QY 116 totcccacagagctggaatggtgatgaattctttttaatcaaaagacaattcttttc 175  
 133 TCTCCGACAGAGCTGGAATG-----C 155  
 QY 176 attgacatttgactgaaacagagcgtattgatgctataataacgtgtgacaatgagc 235  
 DB 156 ATTGACTTTGACTATGGAACACAGAGCTATTGATGCTATATACGTTGACATGAGC 215  
 QY 236 ttaccaccgaagggagcactccaataggcaattgacacctcaacctcaagcacaccatg 295  
 DB 216 TTTACCCGAAGGAGCAGCTCAATATATGGCAATTGACCTCACTCAAGCACACCCCAATG 275  
 QY 296 gacagctgctcaccagctacatgacacagacagattcagtaagctagaagaatgaga 355  
 DB 276 GACAGCTGCTCACCAGTACATGACACGACAAATTCAGTAAAGCTAGAAATGACGA 335  
 QY 356 gtgatgaagagtgtgacagaacccctgagccgtgaagatgagatgagggccatgatg 415  
 DB 336 GTGATGAAGAGTGTGACAGAAACCCCTGAGCCGTGAAATGAGATCAGGGGCCATGATG 395  
 QY 416 aggtatcaccttagaagaacccctaatgagacagcagaggtggctgacaacagggaag 475  
 DB 396 AGGTAGACCCCTAGAGAACCCCTTAATTGAGACGACGAGGTGGCTGACACAGGAAG 455  
 QY 476 tccaggagcttcaagcggaggaggaatccggcttcggaatggttaaaactgaaatgtgacg 535  
 DB 456 TCCAGGAGCTTCAAGCGAGGGAGGAATCCGGCTTCGGAATGTTAAACTGAAATGTGACG 515

QY 536 tctgtggcattgttgcattggggcccaatgtgcttattggtacataaaaggagtcacactg 595  
 DB 516 TCTGTGGCATTGGTTGGTTCATTTGGGCCCAATGTGCTATGGTACATAAAGAGTCACTAG 575  
 QY 596 gtgaacgccccctccactgtaacccagtgtggagcttcttttactcagaaggcgaaccttc 555  
 DB 576 GTGAACGCCCTTCCACTGTAAACAGTGTGGAGCTTCTTTTACTCAGAAGGCAACCTTC 635  
 QY 656 tgagacacataaaagttaacactctggagagaagccgttcaaaatgtccctttctgtagtacag 715  
 DB 636 TGAGACACATAAAGTTACACTCTGGAGAGAAAGCCGTTCAAAATGTCCTTCTGTAGTAGC 695  
 QY 716 cctgtagaagaaggagccctccacaggatcactcaggacccttctgtgggttaaacctc 775  
 DB 696 CCGTGTAGAAGAGGAGCGCCCTCACAGGACACCTCAGAGCCCAATTCGTGGGTAAACCTC 755  
 QY 776 acaagtgaactactgtggacgaagctacaagcagcgcagttcactcggaggagcacaagg 835  
 DB 756 ACAAGTGAACACTCTGTGGAGCAAGCTCAAGCAGCGCAGTTCACTGGAGGAGCACAAAG 815  
 QY 836 aacggtgcacaaactctccagaatgtcagcatgtggagctggaggtcgtgggcagggtcagtc 895  
 DB 816 AACGCTGCCCAACTATCTCCAGATGTTCAGCATGGAGGCTGCTGGGCAGGTCATGAGTC 875  
 QY 896 accatgtacctctctatggaagattgtaagaaacaagagcctattattggacaacaatttt 955  
 DB 876 ACCATGTACTCTCTATGGAAGATTGTAAGAACAGAGCCTATTATGGACAACAATATT 935  
 QY 956 ctctggtgcttttgagagacctgctgctcatagagaagctcagggggaatatgggaaac 1015  
 DB 936 CTCTGCTGCTTTTGAGAGACCTGCTGTCTATAGAAAGCTCACGGGGAATATGGGAAAC 995  
 QY 1016 gtaaaagctccactcccaaaaattgttggggaaagctcagctcagctcagctacccag 1075  
 DB 996 GTAAAGCTCCACTCCACAAAAGTTTGTGGGGAAAAGCTCATGCGATTTCAGTACCCAG 1055  
 QY 1076 atattcaatttgatgaacttaacatatgagaaggaggtgagctgagctgagctcagctc 1135  
 DB 1056 ATATTCACTTTGATGAACTTAACATATGAGAAGGAGGCTGAGCTGATGCTCATATA 1115  
 QY 1136 tgatggacaaagccatcaacaatgcaatcacctaccttggagctgagggcccttcaacctc 1195  
 DB 1116 TGATGGACCAAGCCATCAACAATGAATCACTACCTACCTTGGAGCTGAGGCCCTTCACCTC 1175  
 QY 1196 tgatgacagcccccccaagcaaatcgtcgaagtggccccagttataagctcagctatt 1255  
 DB 1176 TGATGACAGCCCCCCCAGCAAACTGCTGAAGTGGCCCCAGTTATAAGCTCAGCTTATT 1235  
 QY 1256 ctcaaggtctatcccaaatagatagaagacccttagcagggaaactgctgtagtc 1315  
 DB 1236 CTCAGGCTATATCAATCCAAATAGGATAGAAAGACCCATTAGCAGGGAACCTGCTGATAGTC 1295  
 QY 1316 atgaaaaacaatggatggccccatctctctcatcagacaaagagtcgacccccggaaa 1375  
 DB 1296 ATGAAAAACAATGGATGGCCCCATCTCTCATAGACCAAAAGAGTCGACCCCAAGAAA 1355  
 QY 1376 gagagcctctccacagcaatagctgctggatccactgactcagaaagcagccatgatg 1435  
 DB 1356 GAGAGGCTCTCCCAAGCAATAGTGCCTGGATTCCACTGACTCAGAAAGCAGCCATGATG 1415  
 QY 1436 accacagctctctacccaaggacacctgccttaattcccaagagagaaacaaagccagctt 1495  
 DB 1416 ACCACAGCTCTCAACAGGACACCTGCTTAAATCCCAAGAGAAACAACAGCCAGCTT 1475  
 QY 1496 acatgaaggagatgtcaaaagcttggatactacccaaggtccttaaggggtctctgaagg 1555  
 DB 1476 ACATGAAGAGGATGTCAAAGCTTTGGATACTTACCAGGCTCTTACAGGCTCTCTGAGG 1535  
 QY 1556 acatctacaaggtcttcaatggggaaaggagaaacagattagggtccttaaggtgtgagcact 1615  
 DB 1536 ACATCTACAAGGTCTTCAATGGAGAAAGGAGAACAGATTAGGGCTTCAAGTGTGAGCACT 1595  
 QY 1616 gcgaggtctcttctctagaccatgctcatgtacaccattcacatgggtgttgccatgggtacc 1675



